Available online at www.sciencedirect.com

SCIENCE DIRECT

Theoretical Computer Science 360 (2006) 373–385

Theoretical Computer Science

www.elsevier.com/locate/tcs

Approximating and computing behavioural distances in probabilistic transition systems

Franck van Breugel[a,*,1], James Worrell[b]

[a]York University, Department of Computer Science, 4700 Keele Street, Toronto, Canada M3J 1P3

[b]Oxford University, Computing Laboratory, Parks Road, Oxford, OX1 3QD, UK

Received 12 May 2005; received in revised form 28 October 2005; accepted 11 May 2006

Communicated by M.W. Mislove

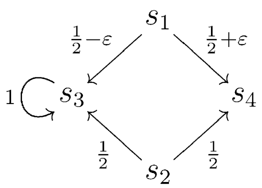
Abstract

In an earlier paper we presented a pseudometric on the states of a probabilistic transition system, yielding a quantitative notion of behavioural equivalence. The behavioural pseudometric was defined via the terminal coalgebra of a functor based on a metric on Borel probability measures. In the present paper we give a polynomial-time algorithm, based on linear programming, to calculate the distances between states up to a prescribed degree of accuracy.

Keywords: Probabilistic transition system; Behavioural pseudometric; Approximation; Coalgebra; Linear programming

1. Introduction

In this paper, we consider reactive probabilistic transition systems. One of the standard equivalences for such systems is *probabilistic bisimulation*, introduced by Larsen and Skou [27]. Briefly, a probabilistic bisimulation is an equivalence relation on states such that for any two related states their probability of making a transition to any equivalence class is equal. Two states are either bisimilar or they are not bisimilar, and a slight change in the probabilities associated to a system can cause bisimilar states to become nonbisimilar and vice-versa. Consider, for example, the system depicted in the following diagram.



* Corresponding author. Tel.: +1 416 736 2100.
E-mail address: franck@cs.yorku.ca (F. van Breugel).

[1] Supported by Natural Sciences and Engineering Research Council of Canada.

0304-3975/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.tcs.2006.05.021

The states s_1 and s_2 are only bisimilar if ε is 0. However, the states give rise to almost the same behaviour for very small ε different from 0.

Motivated by such examples, Giacalone et al. define in [17] a *pseudometric* on the states of a (restricted type of) probabilistic transition system. This yields a *quantitative* notion of behavioural equivalence. A pseudometric differs from an ordinary metric in that different elements, that is, states, can have distance 0. The distance between states, a real number between 0 and 1, is used to express the similarity of the behaviour of those states. The smaller the distance, the more alike the behaviour. In particular, the distance between states is 0 if they are behaviourally indistinguishable.

In [5,7], we presented a *behavioural pseudometric* for reactive probabilistic transition systems. In fact, we introduced a family of pseudometrics, parametric in a constant c in the interval $(0, 1)$—we will discuss the significance of this constant later. For instance, the distance between states s_1 and s_2 in the above example is $c \cdot \varepsilon$. Also, 0-distance coincides with probabilistic bisimilarity.

It turns out that our pseudometric coincides (modulo some minor details) with the one presented by Desharnais et al. [12,13,15]. Their pseudometric is defined by giving a real-valued semantics to a probabilistic modal logic where, in particular, the modal connective is interpreted by integration. The proof that the two pseudometrics coincide (see [5,7]) can be seen as a quantitative analogue of the logical characterization of bisimulation [20]. It is our coinductive presentation of the pseudometric that we will exploit in this paper.

The main contribution of the present paper is a polynomial-time *algorithm* to *approximate* distances in our pseudometric up to a prescribed degree of accuracy. As we explain below, the key ingredients of our pseudometric and algorithm are coalgebras, a metric on Borel probability measures and linear programming.

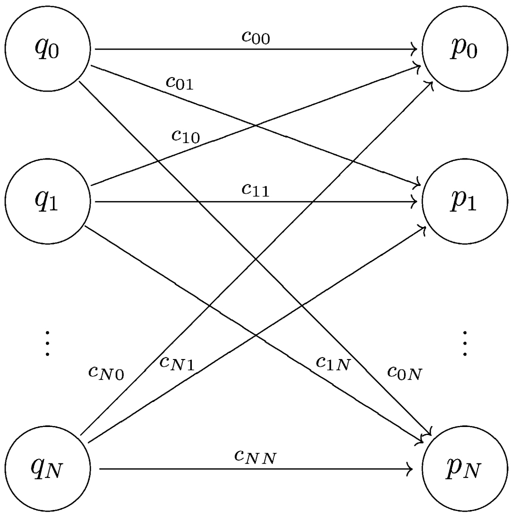
Many different kinds of transition system can be viewed as *coalgebras*; Rutten [33] provides numerous examples. Given an endofunctor $\mathcal{F}$, an $\mathcal{F}$-coalgebra consists of an object S, which is called the carrier, and an arrow $t : S \to \mathcal{F}(S)$. The object S represents the state space and the arrow t captures the transitions. de Vink and Rutten [37] have shown that probabilistic transition systems can be represented as $\mathcal{P}'$-coalgebras, where $\mathcal{P}'$ is an endofunctor on the category of 1-bounded complete ultrametric spaces and nonexpansive functions. Furthermore, they have proved that the functor $\mathcal{P}'$ is locally contractive (a natural generalization of contractive function to the setting of functors). Hence, according to Rutten and Turi's (ultra)metric terminal coalgebra theorem [35], there exists a *terminal* $\mathcal{P}'$-coalgebra. A terminal coalgebra can be viewed as a universal system as it minimally realises all behaviour. By definition, there is a unique map from the carrier of an arbitrary coalgebra to the carrier of the terminal coalgebra. This map preserves and reflects transitions. de Vink and Rutten have also shown that the *kernel* of this unique map from a $\mathcal{P}'$-coalgebra to the terminal $\mathcal{P}'$-coalgebra is probabilistic bisimilarity. That is, states are mapped to the same element in the terminal $\mathcal{P}'$-coalgebra by the unique map if and only if they are probabilistic bisimilar.

In this paper, we study a variation on the endofunctor $\mathcal{P}'$. Our endofunctor $\mathcal{P}$ on the category $\mathbb{CM}et_1$ of 1-bounded complete metric spaces and nonexpansive functions is based on a *metric* on Borel probability measures. This metric arises in very different contexts and under many different names, including the Hutchinson metric [21], the Kantorovich metric [24] and the Vaserstein metric [36]. Like $\mathcal{P}'$-coalgebras, $\mathcal{P}$-coalgebras can also represent probabilistic transition systems, as we will show in Example 15. Furthermore, we will prove that the functor $\mathcal{P}$ is locally contractive and, hence, has a terminal coalgebra. Since the terminal $\mathcal{P}$-coalgebra carries a metric, we can also consider the *pseudometric kernel* of the unique map ϕ from a $\mathcal{P}$-coalgebra to the terminal $\mathcal{P}$-coalgebra. This is a pseudometric on the carrier of the $\mathcal{P}$-coalgebra. The distance between states s_1 and s_2 of a $\mathcal{P}$-coalgebra, that is, a probabilistic transition system, is the distance in the terminal $\mathcal{P}$-coalgebra of their images $\phi(s_1)$ and $\phi(s_2)$. Since our functor is similar to the one considered by de Vink and Rutten, we still have that states are bisimilar if and only if they are mapped to the same element in the terminal $\mathcal{P}$-coalgebra and, hence, have distance 0.

As Rutten and Turi [35] have shown, the unique map from an $\mathcal{F}$-coalgebra to the terminal $\mathcal{F}$-coalgebra, where $\mathcal{F}$ is a locally contractive endofunctor on the category $\mathbb{CM}et_1$, can be defined as the unique fixed point $\mathit{fix}(\Phi)$ of a function Φ from a nonempty complete metric space to itself. Since the functor $\mathcal{F}$ is locally contractive, the function Φ is contractive. Hence, according to Banach's fixed point theorem [2], Φ has a unique fixed point $\mathit{fix}(\Phi)$. This fixed point can be approximated by a sequence of functions $(\phi_n)_n$. The function ϕ_0 is an arbitrary constant function from the $\mathcal{F}$-coalgebra to the terminal $\mathcal{F}$-coalgebra and the other functions are defined by $\phi_n = \Phi(\phi_{n-1})$. The pseudometric kernel of the unique map $\mathit{fix}(\Phi)$ assigns distances to the elements of the carrier of the $\mathcal{F}$-coalgebra. We denote this pseudometric by $d_{\mathit{fix}(\Phi)}$. The approximations ϕ_n of $\mathit{fix}(\Phi)$ also induce pseudometrics on the carrier of the $\mathcal{F}$-coalgebra: the pseudometric kernels of ϕ_n. We denote these pseudometrics by d_{ϕ_n}. We will show that the pseudometric $d_{\mathit{fix}(\Phi)}$

can be approximated by the pseudometrics d_{ϕ_n}. In particular, to calculate the $d_{fix(\Phi)}$-distances to a prescribed degree of accuracy α, we only have to calculate the $\phi_1, \ldots, \phi_{\log_c(\alpha/2)}$-distances.

As we will see, the problem of computing ϕ_{n+1} from ϕ_n can be reduced to a *linear programming* problem. In particular, it can be reduced to more specific problems including the *minimum cost network flow* problem and the *minimum cost circulation* problem. These problems can be visualized as follows.



The aim is to minimize the cost of transportation from the origins on the left to the destinations on the right. Each origin node is labelled with its supply and each destination node with its demand. Furthermore, each origin is connected to each destination via an edge labelled with its cost. For a detailed discussion of such a transportation problem we refer the reader to, for example, [8, Chapter 21]. Problems like the minimum network flow problem [30] and the minimum cost circulation problem [34] can be solved in polynomial time.

2. Probabilistic transition systems and probabilistic bisimulation

We present the probabilistic analogues of labelled transition systems and bisimulation. For simplicity we only consider unlabelled transitions, though all our results generalize to the labelled case.

Definition 1. A *probabilistic transition system* consists of a finite set S of states together with a transition function $\pi : S \times S \to [0, 1]$ such that, for each $s \in S$, $\sum_{s' \in S} \pi(s, s') \leqslant 1$.

The transition function π is a conditional sub-probability distribution. It determines the interaction of the system with the environment. $\pi(s, s')$ is the probability that the system ends up in state s' given that it was in state s before the interaction. This interpretation of π is known as the reactive model and it is due to Larsen and Skou [27].

We impose the restriction $\sum_{s' \in S} \pi(s, s') \leqslant 1$ instead of the more common, but also more restrictive, condition $\sum_{s' \in S} \pi(s, s') \in \{0, 1\}$. We interpret $1 - \sum_{s' \in S} \pi(s, s')$ as the probability that the system refuses to interact with the environment when the system is in state s.

Larsen and Skou [27] adapted bisimulation—a key behavioural equivalence for labelled transition systems due to Milner [29] and Park [31]—for probabilistic transition systems as follows.

Definition 2. Let $\langle S, \pi \rangle$ be a probabilistic transition system. An equivalence relation $\mathcal{R}$ on the set of states S is a *probabilistic bisimulation* if $s_1 \mathcal{R} s_2$ implies $\sum_{s' \in E} \pi(s_1, s') = \sum_{s' \in E} \pi(s_2, s')$ for all $\mathcal{R}$-equivalence classes E. States s_1 and s_2 are *probabilistic bisimilar* if $s_1 \mathcal{R} s_2$ for some probabilistic bisimulation $\mathcal{R}$.

Example 3. Consider the probabilistic transition system presented in the introduction and assume that ε equals 0. The smallest equivalence relation containing (s_1, s_2) is a probabilistic bisimulation. Hence, the states s_1 and s_2 are probabilistic bisimilar.

3. A metric terminal coalgebra theorem

We introduce coalgebras and Rutten and Turi's metric terminal coalgebra theorem [35]. For more details about the theory of coalgebra we refer the reader to, for example, [22].

Definition 4. Let $\mathbb{C}$ be a category. Let $\mathcal{F} : \mathbb{C} \to \mathbb{C}$ be a functor. An *$\mathcal{F}$-coalgebra* consists of an object C in $\mathbb{C}$ together with an arrow $f : C \to \mathcal{F}(C)$ in $\mathbb{C}$. The object C is called the *carrier*. An *$\mathcal{F}$-homomorphism* from $\mathcal{F}$-coalgebra $\langle C, f\rangle$ to $\mathcal{F}$-coalgebra $\langle D, g\rangle$ is an arrow $\phi : C \to D$ in $\mathbb{C}$ such that $\mathcal{F}(\phi) \circ f = g \circ \phi$.

$$\begin{tikzcd} C \arrow[r, "\phi"] \arrow[d, "f"'] & D \arrow[d, "g"] \\ \mathcal{F}(C) \arrow[r, "\mathcal{F}(\phi)"'] & \mathcal{F}(D) \end{tikzcd}$$

The $\mathcal{F}$-coalgebras and $\mathcal{F}$-homomorphisms form a category. If this category has a terminal object, then this object is called the *terminal $\mathcal{F}$-coalgebra*.

We restrict our attention to the category $\mathbb{CM}et_1$ of 1-bounded complete metric spaces and nonexpansive functions. A metric space X is 1-bounded if all its distances are bounded by 1, that is, for all $x_1, x_2 \in X$, $d_X(x_1, x_2) \leqslant 1$. A function $f : X \to Y$ is nonexpansive if it does not increase any distances, that is, for all $x_1, x_2 \in X$, $d_Y(f(x_1), f(x_2)) \leqslant d_X(x_1, x_2)$. We denote the collection of nonexpansive functions from the space X to the space Y by $X \underset{1}{\to} Y$. This collection can be turned into a metric space by endowing the functions with the supremum metric, that is, for $f_1, f_2 \in X \underset{1}{\to} Y$, $d_{X \underset{1}{\to} Y}(f_1, f_2) = \sup_{x \in X} d_Y(f_1(x), f_2(x))$.

Let c be a constant in the interval $(0, 1)$. A function $f : X \to Y$ is c-contractive if it decreases all distances by at least a factor c, that is, for all $x_1, x_2 \in X$, $d_Y(f(x_1), f(x_2)) \leqslant c \cdot d_X(x_1, x_2)$.

Definition 5. A functor $\mathcal{F} : \mathbb{CM}et_1 \to \mathbb{CM}et_1$ is *locally c-contractive* (*locally nonexpansive*) if for all 1-bounded complete metric spaces X and Y, the function $\mathcal{F}_{X,Y} : (X \underset{1}{\to} Y) \to (\mathcal{F}(X) \underset{1}{\to} \mathcal{F}(Y))$ defined by

$$\mathcal{F}_{X,Y}(f) = \mathcal{F}(f)$$

is c-contractive (nonexpansive).

In the rest of this section, we restrict ourselves to locally c-contractive functors. For these functors, we have

Theorem 6 (*Turi and Rutten [35, Theorem 4.8]*). *There exists a terminal $\mathcal{F}$-coalgebra $\langle fix(\mathcal{F}), \iota\rangle$.*

From (the dual of) Lambek's lemma [26, Example 0] we can conclude that ι is an isomorphism and, hence, its inverse ι^{-1} exists and is isometric. Recall that a function $f : X \to Y$ is isometric if it preserves all distances, that is, for all $x_1, x_2 \in X$, $d_Y(f(x_1), f(x_2)) = d_X(x_1, x_2)$.

For the rest of this section, we fix $\langle X, \mu\rangle$ to be an $\mathcal{F}$-coalgebra. To characterize the unique map from the $\mathcal{F}$-coalgebra $\langle X, \mu\rangle$ to the terminal $\mathcal{F}$-coalgebra $\langle fix(\mathcal{F}), \iota\rangle$ we introduce the following function.

Definition 7. The function $\Phi : (X \underset{1}{\to} \mathit{fix}(\mathcal{F})) \to (X \underset{1}{\to} \mathit{fix}(\mathcal{F}))$ is defined by $\Phi(\phi) = \iota^{-1} \circ \mathcal{F}(\phi) \circ \mu$.

$$\begin{tikzcd} X \arrow[r, "\phi"] \arrow[d, "\mu"'] & \mathit{fix}(\mathcal{F}) \\ \mathcal{F}(X) \arrow[r, "\mathcal{F}(\phi)"'] & \mathcal{F}(\mathit{fix}(\mathcal{F})) \arrow[u, "\iota^{-1}"'] \end{tikzcd}$$

Note that the above diagram does not commute.

Since the functor $\mathcal{F}$ is locally c-contractive, the function Φ is c-contractive (see proof of [35, Theorem 4.5]). Since Φ is a contractive function from a nonempty complete metric space to itself, we can conclude from Banach's fixed point theorem [2, Theorem II.6] that it has a unique fixed point $\mathit{fix}(\Phi)$. The function $\mathit{fix}(\Phi)$ is the unique $\mathcal{F}$-homomorphism from the $\mathcal{F}$-coalgebra $\langle X, \mu\rangle$ to the terminal $\mathcal{F}$-coalgebra $\langle \mathit{fix}(\mathcal{F}), \iota\rangle$ (see proof of [35, Theorem 4.5]). We conclude this section by showing that the unique map $\mathit{fix}(\Phi)$ can be approximated by the maps ϕ_n.

Definition 8. Let $\phi_0 : X \underset{1}{\to} \mathit{fix}(\mathcal{F})$ be some constant function. For $n > 0$, the function $\phi_n : X \underset{1}{\to} \mathit{fix}(\mathcal{F})$ is defined by

$$\phi_n = \Phi(\phi_{n-1}).$$

Proposition 9. *For all* $n \geqslant 0$,

$$d_{X \underset{1}{\to} \mathit{fix}(\mathcal{F})}(\phi_n, \mathit{fix}(\Phi)) \leqslant c^n.$$

Proof. By induction on n. Obviously, the property holds for $n = 0$. Let $n > 0$. Then

$$\begin{aligned} & d_{X \underset{1}{\to} \mathit{fix}(\mathcal{F})}(\phi_n, \mathit{fix}(\Phi)) \\ & \quad = d_{X \underset{1}{\to} \mathit{fix}(\mathcal{F})}(\Phi(\phi_{n-1}), \Phi(\mathit{fix}(\Phi))) \\ & \quad \leqslant c \cdot d_{X \underset{1}{\to} \mathit{fix}(\mathcal{F})}(\phi_{n-1}, \mathit{fix}(\Phi)) \quad [\Phi \text{ is } c\text{-contractive}] \\ & \quad \leqslant c \cdot c^{n-1} \quad [\text{induction}] \\ & \quad = c^n. \qquad \square \end{aligned}$$

4. Pseudometric kernels

Our behavioural pseudometric on the states of a probabilistic transition system will be defined as the so-called pseudometric kernel induced by the unique map from the probabilistic transition system, viewed as a coalgebra, to the terminal coalgebra. Below, we introduce pseudometric kernels. Furthermore, we show that the pseudometric kernel induced by $\mathit{fix}(\Phi)$ can be approximated by the pseudometric kernels induced by ϕ_n.

A function ϕ from the space X to the space $\mathit{fix}(\mathcal{F})$ defines a distance function d_ϕ on X. We call this distance function the pseudometric kernel induced by ϕ. The distance between x_1 and x_2 in X is defined as the distance of their ϕ-images in the metric space $\mathit{fix}(\mathcal{F})$.

Definition 10. Let $\phi \in X \to \mathit{fix}(\mathcal{F})$. The distance function $d_\phi : X \times X \to [0, 1]$ is defined by

$$d_\phi(x_1, x_2) = d_{\mathit{fix}(\mathcal{F})}(\phi(x_1), \phi(x_2)).$$

One can easily verify that the distance function d_ϕ is a pseudometric. Recall that different elements may have distance 0 in a pseudometric space. Note that the elements x_1 and x_2 of the pseudometric space $\langle X, d_\phi\rangle$ have distance 0 only if they are mapped by ϕ to the same element in $\mathit{fix}(\mathcal{F})$, since $\mathit{fix}(\mathcal{F})$ is a metric space.

The pseudometric $d_{\mathit{fix}(\Phi)}$ can be approximated by the pseudometrics d_{ϕ_n} as is shown in

Proposition 11. *For all* $n \geqslant 0$ *and* $x_1, x_2 \in X$,

$$|d_{\phi_n}(x_1, x_2) - d_{\mathit{fix}(\Phi)}(x_1, x_2)| \leqslant 2 \cdot c^n.$$

Proof.

$$\begin{aligned}
&|d_{\phi_n}(x_1, x_2) - d_{\mathit{fix}(\Phi)}(x_1, x_2)| = |d_{\mathit{fix}(\mathcal{F})}(\phi_n(x_1), \phi_n(x_2)) - d_{\mathit{fix}(\mathcal{F})}(\mathit{fix}(\Phi)(x_1), \mathit{fix}(\Phi)(x_2))| \\
&\quad \leqslant d_{\mathit{fix}(\mathcal{F})}(\phi_n(x_1), \mathit{fix}(\Phi)(x_1)) + d_{\mathit{fix}(\mathcal{F})}(\phi_n(x_2), \mathit{fix}(\Phi)(x_2)) \quad \text{[triangle inequality]} \\
&\quad \leqslant 2 \cdot d_{X \underset{1}{\rightarrow} \mathit{fix}(\mathcal{F})}(\phi_n, \mathit{fix}(\Phi)) \\
&\quad \leqslant 2 \cdot c^n \quad \text{[Proposition 9]}. \qquad \square
\end{aligned}$$

To compute the $d_{\mathit{fix}(\Phi)}$-distances up to accuracy α, it suffices to calculate the $d_{\phi_{\lceil \log_c(\alpha/2) \rceil}}$-distances.

Proposition 12. *For all* $0 < \alpha < 1$ *and* $x_1, x_2 \in X$,

$$|d_{\phi_{\lceil \log_c(\alpha/2) \rceil}}(x_1, x_2) - d_{\mathit{fix}(\Phi)}(x_1, x_2)| \leqslant \alpha.$$

Proof.

$$\begin{aligned}
&|d_{\phi_{\lceil \log_c(\alpha/2) \rceil}}(x_1, x_2) - d_{\mathit{fix}(\Phi)}(x_1, x_2)| \\
&\quad \leqslant 2 \cdot c^{\lceil \log_c(\alpha/2) \rceil} \quad \text{[Proposition 11]} \\
&\quad \leqslant 2 \cdot c^{\log_c(\alpha/2)} \\
&\quad = \alpha. \qquad \square
\end{aligned}$$

5. A metric on Borel probability measures

We present a metric on the set of Borel probability measures on a metric space. We restrict ourselves to 1-bounded complete metric spaces. We also only consider tight Borel probability measures. A measure μ is tight if it is completely determined by its values for the compact subsets of the metric space X, that is, for all $\varepsilon > 0$ there exists a compact subset K_ε of X such that $\mu(X \setminus K_\varepsilon) < \varepsilon$. We denote the set of tight Borel probability measures on X by $\mathcal{M}(X)$. Under quite mild conditions on the space, for example, compactness, every measure is tight (see, for example, [32, Section II.3]). In particular, all probabilistic transition systems can be represented using tight measures as we will see in Example 15. A distance function on $\mathcal{M}(X)$ is introduced in the following definition.

Definition 13. The function $d_{\mathcal{M}(X)} : \mathcal{M}(X) \times \mathcal{M}(X) \to [0, 1]$ is defined by

$$d_{\mathcal{M}(X)}(\mu_1, \mu_2) = \sup \left\{ \int_X f \, \mathrm{d}\mu_1 - \int_X f \, \mathrm{d}\mu_2 \mid f \in X \underset{1}{\rightarrow} [0, \infty) \right\}.$$

For a proof that $d_{\mathcal{M}(X)}$ is a 1-bounded complete metric, we refer the reader to, for example, [16, Section 2.5].

In [7, Section 4] we show that $\mathcal{M}$ can be extended to an endofunctor on the category $\mathbb{CM}et_1$ by defining $\mathcal{M}(f)$: $\mathcal{M}(X) \to \mathcal{M}(Y)$ by $\mathcal{M}(f)(\mu) = \mu \circ f^{-1}$ for $f \in X \underset{1}{\rightarrow} Y$.

Later, we will exploit the following proposition.

Proposition 14. *Let* $f \in X \underset{1}{\rightarrow} Y$, $g \in Y \underset{1}{\rightarrow} [0, \infty)$ *and* $\mu \in \mathcal{M}(X)$. *Then* $\int_Y g \, \mathrm{d}(\mathcal{M}(f)(\mu)) = \int_X (g \circ f) \, \mathrm{d}\mu$.

Proof. See, for example, the proof of [18, Theorem 1]. $\square$

6. The functor $\mathcal{P}$

We show that probabilistic transition systems can be seen as coalgebras of an endofunctor $\mathcal{P}$ on $\mathbb{CM}et_1$. As we will see, this functor is locally contractive. Hence, by Theorem 6, it has a terminal coalgebra. The pseudometric kernel of the unique map to the terminal $\mathcal{P}$-coalgebra defines a pseudometric on the carrier of an arbitrary $\mathcal{P}$-coalgebra and, hence, on the states of a probabilistic transition system.

The functor $\mathcal{M}$, which we introduced in the previous section, is the key ingredient of our functor $\mathcal{P}$. The following objects also play a role in the definition of $\mathcal{P}$:

- The terminal object of $\mathbb{CM}et_1$ is the singleton space $\mathbf{1}$ whose sole element we denote by $\mathbf{0}$.
- The coproduct object of the objects X and Y in $\mathbb{CM}et_1$ is the disjoint union of the sets underlying the spaces X and Y endowed with the metric

$$d_{X+Y}(v, w) = \begin{cases} d_X(v, w) & \text{if } v \in X \text{ and } w \in X, \\ d_Y(v, w) & \text{if } v \in Y \text{ and } w \in Y, \\ 1 & \text{otherwise.} \end{cases}$$

- Let $c \in (0, 1)$. The scaling by $c\,\cdot$ of an object in $\mathbb{CM}et_1$ leaves the set unchanged and multiplies all distances by c.

Now, we are ready to present the functor $\mathcal{P}$. But first we introduce the functor $\mathcal{R}$ which models refusal:

$$\mathcal{R} = \mathbf{1} + c \cdot - : \mathbb{CM}et_1 \to \mathbb{CM}et_1.$$

The functor $\mathcal{P}$ is defined by

$$\mathcal{P} = \mathcal{M} \circ \mathcal{R} : \mathbb{CM}et_1 \to \mathbb{CM}et_1.$$

Every probabilistic transition system can be seen as a $\mathcal{P}$-coalgebra as is demonstrated in

Example 15. Let $\langle S, \pi\rangle$ be a probabilistic transition system. We endow the set of states S with the discrete metric, that is, for all $s_1, s_2 \in S$,

$$d_S(s_1, s_2) = \begin{cases} 0 & \text{if } s_1 = s_2, \\ 1 & \text{otherwise.} \end{cases}$$

Consequently, every subset of the 1-bounded complete metric space $\mathcal{R}(S)$ is a Borel set. For every state s, the Borel probability measure μ_s is the discrete Borel probability measure determined by

$$\begin{aligned} &\mu_s(\mathbf{1}) = 1 - \textstyle\sum_{s' \in S} \pi(s, s'), \\ &\mu_s(\{s'\}) = \pi(s, s'). \end{aligned}$$

Since the space S is finite, $\mathcal{R}(S)$ is finite and, hence, compact. Therefore, the measure μ_s is tight. Because S is endowed with the discrete metric, the function μ mapping the state s to the measure μ_s is nonexpansive. Hence, every probabilistic transition system can be viewed as a $\mathcal{P}$-coalgebra.

$\mathcal{P}$-coalgebras can also capture probabilistic systems the state space of which is continuous. We refer the reader to [7] for examples.

As shown by America and Rutten in [1, Theorem 5.4], the functor $c\,\cdot$ is locally c-contractive and the functor $+$ is locally nonexpansive. In [7, Proposition 17], we show that the functor $\mathcal{M}$ is locally nonexpansive as well. From these facts and [1, Theorem 5.4] we can conclude that the functor $\mathcal{P}$ is locally c-contractive. Thus, according to Theorem 6, there exists a terminal $\mathcal{P}$-coalgebra. Our behavioural pseudometric on a probabilistic transition system is defined as the pseudometric kernel $d_{\mathit{fix}(\Phi)}$ where $\mathit{fix}(\Phi)$ is the unique map from the probabilistic transition system, viewed as a $\mathcal{P}$-coalgebra, to the terminal $\mathcal{P}$-coalgebra. In this pseudometric, states have distance 0 only if they are probabilistic bisimilar.

Proposition 16 (*van Breugel and Worrell [7, Proposition 32]*). *States s_1 and s_2 are probabilistic bisimilar if and only if $d_{\mathit{fix}(\Phi)}(s_1, s_2) = 0$.*

More generally, the distance between states is a trade-off between the depth of observations needed to distinguish the states and the amount each observation differentiates the states. The relative weight given to these two factors is determined by the discount factor c lying between 0 and 1: the smaller the value of c the greater the discount on observations made at greater depth.

Example 17. *For the system depicted in the introduction, the distances are given in the table below*:

	s_1	s_2	s_3	s_4
s_1	0			
s_2	$\varepsilon \cdot c$	0		
s_3	$(\frac{1}{2} + \varepsilon) \cdot c$	$\frac{1}{2} \cdot c$	0	
s_4	1	1	1	0

7. Pseudometric kernels for $\mathcal{P}$-coalgebras

Our behavioural pseudometric on a probabilistic transition system is defined as the pseudometric kernel $d_{\mathit{fix}(\Phi)}$ where $\mathit{fix}(\Phi)$ is the unique map from the probabilistic transition system, viewed as a $\mathcal{P}$-coalgebra, to the terminal $\mathcal{P}$-coalgebra. As we have already seen in Section 4, $d_{\mathit{fix}(\Phi)}$ can be approximated by the pseudometric kernels d_{ϕ_n}. Next, we present a characterization of the pseudometrics d_{ϕ_n} for $\mathcal{P}$-coalgebras. Furthermore, we show that the d_{ϕ_n}-distances are smaller than or equal to the $d_{\mathit{fix}(\Phi)}$-distances.

To prove our characterizations, we use the following:

Proposition 18. *Let* $\phi \in X \to \mathit{fix}(\mathcal{P})$. *Then*

$$\mathcal{R}\langle X, d_\phi \rangle \underset{1}{\to} [0, \infty) = \{\, g \circ \mathcal{R}(\phi) \mid g \in \mathcal{R}(\mathit{fix}(\mathcal{P})) \underset{1}{\to} [0, \infty)\}.$$

Proof. By a slight abuse of notation, ϕ may be regarded as an isometric embedding of the pseudometric space $\langle X, d_\phi \rangle$ in $\mathit{fix}(\mathcal{P})$. Thus $\mathcal{R}(\phi)$ is an isometric embedding of $\mathcal{R}\langle X, d_\phi \rangle$ into $\mathcal{R}(\mathit{fix}(\mathcal{P}))$.

According to the McShane–Whitney extension lemma [28, Theorem 1] and [38, footnote on p. 63], for every isometric embedding $i : X \to Y$ and nonexpansive function $f : X \to Z$, there exists a nonexpansive function $g : Y \to Z$ such that $g \circ i = f$.

$$\begin{tikzcd}
\mathcal{R}\langle X, d_\phi \rangle \arrow[rr, "\mathcal{R}(\phi)"] \arrow[dr, "f"'] & & \mathcal{R}(\mathit{fix}(\mathcal{P})) \arrow[dl, "g"] \\
& {[0, \infty)} &
\end{tikzcd}$$

Hence, every nonexpansive map $f : \mathcal{R}\langle X, d_\phi \rangle \to [0, \infty)$ has a nonexpansive extension $g : \mathcal{R}(\mathit{fix}(\mathcal{P})) \to [0, \infty)$ in the sense that $g \circ \mathcal{R}(\phi) = f$.

The other inclusion is trivial. $\square$

We can characterize the pseudometric d_{ϕ_n} on the carrier of a $\mathcal{P}$-coalgebra $\langle X, \mu \rangle$ as follows.

Theorem 19. *For all* $x_1, x_2 \in X$,

$$d_{\phi_0}(x_1, x_2) = 0.$$

For all $n > 0$ and $x_1, x_2 \in X$,

$$d_{\phi_n}(x_1, x_2) = \sup \left\{ \int_{\mathcal{R}(X)} g \, \mathrm{d}\mu_{x_1} - \int_{\mathcal{R}(X)} g \, \mathrm{d}\mu_{x_2} \mid g \in \mathcal{R}\langle X, d_{\phi_{n-1}} \rangle \underset{1}{\to} [0, \infty) \right\}.$$

Proof. Obviously, $d_{\phi_0}(x_1, x_2) = d_{\mathit{fix}(\mathcal{P})}(\phi_0(x_1), \phi_0(x_2)) = 0$, since ϕ_0 is a constant function. Furthermore, for all $n > 0$ we have

$$\begin{aligned}
& d_{\phi_n}(x_1, x_2) \\
&\quad = d_{\mathit{fix}(\mathcal{P})}(\phi_n(x_1), \phi_n(x_2)) \\
&\quad = d_{\mathit{fix}(\mathcal{P})}(\Phi(\phi_{n-1})(x_1), \Phi(\phi_{n-1})(x_2)) \\
&\quad = d_{\mathit{fix}(\mathcal{P})}((\iota^{-1} \circ \mathcal{P}(\phi_{n-1}) \circ \mu)(x_1), (\iota^{-1} \circ \mathcal{P}(\phi_{n-1}) \circ \mu)(x_2)) \quad \text{[Definition 7]} \\
&\quad = d_{\mathcal{P}(\mathit{fix}(\mathcal{P}))}((\mathcal{P}(\phi_{n-1}) \circ \mu)(x_1), (\mathcal{P}(\phi_{n-1}) \circ \mu)(x_2)) \quad [\iota^{-1} \text{ is isometric}] \\
&\quad = \sup \left\{ \int_{\mathcal{R}(\mathit{fix}(\mathcal{P}))} f \, \mathrm{d}((\mathcal{P}(\phi_{n-1}) \circ \mu)(x_1)) \right. \\
&\qquad \left. - \int_{\mathcal{R}(\mathit{fix}(\mathcal{P}))} f \, \mathrm{d}((\mathcal{P}(\phi_{n-1}) \circ \mu)(x_2)) \mid f \in \mathcal{R}(\mathit{fix}(\mathcal{P})) \underset{1}{\to} [0, \infty) \right\} \\
&\quad = \sup \left\{ \int_{\mathcal{R}(X)} (f \circ \mathcal{R}(\phi_{n-1})) \, \mathrm{d}\mu_{x_1} \right. \\
&\qquad \left. - \int_{\mathcal{R}(X)} (f \circ \mathcal{R}(\phi_{n-1})) \, \mathrm{d}\mu_{x_2} \mid f \in \mathcal{R}(\mathit{fix}(\mathcal{P})) \underset{1}{\to} [0, \infty) \right\} \quad \text{[Proposition 14]} \\
&\quad = \sup \left\{ \int_{\mathcal{R}(X)} g \, \mathrm{d}\mu_{x_1} - \int_{\mathcal{R}(X)} g \, \mathrm{d}\mu_{x_2} \mid g \in \mathcal{R}\langle X, d_{\phi_{n-1}} \rangle \underset{1}{\to} [0, \infty) \right\} \quad \text{[Proposition 18].} \qquad \square
\end{aligned}$$

Next, we present a dual representation for d_{ϕ_n}. This representation is based on the Kantorovich–Rubinstein duality theorem [25]. Let Y be a compact metric space. Let v_1 and v_2 be Borel probability measures on Y. (Recall that each Borel probability measure on a compact space is tight.) We denote the set of Borel probability measures on the product space with marginals v_1 and v_2, that is, the Borel probability measures v on Y^2 such that for all Borel subsets B of Y,

$$v(B \times Y) = v_1(B) \quad \text{and} \quad v(Y \times B) = v_2(B),$$

by $v_1 \otimes v_2$. The Kantorovich–Rubinstein duality theorem tells us

$$\sup \left\{ \int_Y f \, \mathrm{d}v_1 - \int_Y f \, \mathrm{d}v_2 \mid f \in Y \underset{1}{\to} [0, \infty) \right\} = \inf \left\{ \int_{Y^2} d_Y \, \mathrm{d}v \mid v \in v_1 \otimes v_2 \right\}.$$

We exploit this theorem in

Corollary 20. *If X is compact then for all $n > 0$ and $x_1, x_2 \in X$,*

$$d_{\phi_n}(x_1, x_2) = \inf \left\{ \int_{\mathcal{R}(X)^2} d_{\mathcal{R}\langle X, d_{\phi_{n-1}} \rangle} \, \mathrm{d}\mu \mid \mu \in \mu_{x_1} \otimes \mu_{x_2} \right\}.$$

Proof.

$$\begin{aligned}
& d_{\phi_n}(x_1, x_2) \\
&\quad = \sup \left\{ \int_{\mathcal{R}(X)} g \, \mathrm{d}\mu_{x_1} - \int_{\mathcal{R}(X)} g \, \mathrm{d}\mu_{x_2} \mid g \in \mathcal{R}\langle X, d_{\phi_{n-1}} \rangle \underset{1}{\to} [0, \infty) \right\} \quad \text{[Theorem 19]} \\
&\quad = \inf \left\{ \int_{\mathcal{R}(X)^2} d_{\mathcal{R}\langle X, d_{\phi_{n-1}} \rangle} \, \mathrm{d}\mu \mid \mu \in \mu_{x_1} \otimes \mu_{x_2} \right\} \quad \text{[Kantorovich–Rubinstein duality theorem].} \qquad \square
\end{aligned}$$

We conclude this section with a proof that the d_{ϕ_n}-distances are smaller than or equal to the $d_{\mathit{fix}(\Phi)}$-distances.

Proposition 21. *For all* $n \geqslant 0$,

$$d_{\phi_n} \leqslant d_{\phi_{n+1}}.$$

Proof. By induction on n. The case $n = 0$ is trivial. Let $n > 0$. For all $x_1, x_2 \in X$,

$$\begin{aligned}
& d_{\phi_n}(x_1, x_2) \\
&= \sup\left\{ \int_{\mathcal{R}(X)} g \, \mathrm{d}\mu_{x_1} - \int_{\mathcal{R}(X)} g \, \mathrm{d}\mu_{x_2} \mid g \in \mathcal{R}\langle X, d_{\phi_{n-1}} \rangle \underset{1}{\rightarrow} [0, \infty) \right\} \quad \text{[Theorem 19]} \\
&\leqslant \sup\left\{ \int_{\mathcal{R}(X)} g \, \mathrm{d}\mu_{x_1} - \int_{\mathcal{R}(X)} g \, \mathrm{d}\mu_{x_2} \mid g \in \mathcal{R}\langle X, d_{\phi_n} \rangle \underset{1}{\rightarrow} [0, \infty) \right\} \\
&\quad \text{[by induction } d_{\phi_{n-1}} \leqslant d_{\phi_n}, \text{ so } \mathcal{R}\langle X, d_{\phi_{n-1}} \rangle \underset{1}{\rightarrow} [0, \infty) \subseteq \mathcal{R}\langle X, d_{\phi_n} \rangle \underset{1}{\rightarrow} [0, \infty)] \\
&= d_{\phi_{n+1}}(x_1, x_2) \quad \text{[Theorem 19].} \qquad \square
\end{aligned}$$

Corollary 22. *For all* $n \geqslant 0$, $d_{\phi_n} \leqslant d_{\mathit{fix}(\Phi)}$.

8. The algorithm

Suppose that the $\mathcal{P}$-coalgebra considered in the previous section represents a probabilistic transition system $\langle S, \pi \rangle$, where $S = \{s_1, \ldots, s_N\}$. Below, we show that the calculation of $d_{\phi_n}(s_k, s_\ell)$ can be reduced to a linear programming problem and that approximating our behavioural pseudometric for $\langle S, \pi \rangle$ up to accuracy α boils down to solving $(N(N-1)/2)\lceil \log_c(\alpha/2) \rceil$ linear programming problems, where each problem has $2N$ nodes and N^2 edges. Since there exist strongly polynomial algorithms to solve this type of linear programming problem (see, for example, [30,34]), our algorithm to approximate the behavioural pseudometric is strongly polynomial as well.

First, we introduce some notation. We add a special state, named s_0, that will play the role of **0**. For $0 \leqslant i, j \leqslant N$ we define $c_{ij} = d_{\mathcal{R}\langle X, d_{\phi_{n-1}} \rangle}(s_i, s_j)$, that is, for $1 \leqslant i, j \leqslant N$, $c_{ij} = c \cdot d_{\phi_{n-1}}(s_i, s_j)$, and $c_{i0} = 1$, $c_{0j} = 1$ and $c_{00} = 0$. Also, we define $p_i = \pi(s_k, s_i)$ and $q_i = \pi(s_\ell, s_i)$ for $1 \leqslant i \leqslant N$. Finally, we define $p_0 = \pi(s_k, s_0) = 1 - \sum_{1 \leqslant i \leqslant N} \pi(s_k, s_i)$ and $q_0 = \pi(s_\ell, s_0) = 1 - \sum_{1 \leqslant i \leqslant N} \pi(s_\ell, s_i)$.

Since integration against discrete measures reduces to summation, according to Corollary 20 to calculate $d_{\phi_n}(s_k, s_\ell)$ we need to

$$\begin{aligned}
\text{minimize} \quad & \sum_{i,j=0}^{N} c_{ij} \cdot \lambda_{ij}, \\
\text{subject to} \quad & \sum_{i=0}^{N} \lambda_{ij} = p_j, \quad 0 \leqslant j \leqslant N, \\
& \sum_{j=0}^{N} \lambda_{ij} = q_i, \quad 0 \leqslant i \leqslant N, \\
& 0 \leqslant \lambda_{ij}, \quad 0 \leqslant i, j \leqslant N.
\end{aligned}$$

The above is a particular type of linear programming problem that we already discussed in the introduction. The network in the introduction corresponds to the above problem.

Now, we are in a position to present our algorithm for calculating the pseudometric $d_{\mathit{fix}(\Phi)}$ on $\langle S, \pi \rangle$ up to a prescribed degree of accuracy α. The algorithm iteratively calculates d_{ϕ_n}, with the value of $d_{\phi_n}(s_k, s_\ell)$ being stored in $d_{k\ell}$. By Proposition 12, $\lceil \log_c(\alpha/2) \rceil$ cycles of the main loop suffice to get within α of $d_{\mathit{fix}(\Phi)}$. Also, recall from Corollary 22 that the d_{ϕ_n} approximate $d_{\mathit{fix}(\Phi)}$ from below.

Step 1 (initialization): We initialize the distance matrix by setting $d_{k\ell} = 0$ for $1 \leqslant k, \ell \leqslant N$. The main body of the algorithm also uses an $N + 1$ by $N + 1$ matrix c, and we initialize some of the entries of this matrix thus: $c_{i0} = 1$ and $c_{0j} = 1$ for $1 \leqslant i, j \leqslant N$, and $c_{00} = 0$. These values never change during the execution of the algorithm.

Step 2 (main loop):
do $\lceil \log_c(\alpha/2) \rceil$ times
for each $1 \leqslant i, j \leqslant N$ do
$c_{ij} = c \cdot d_{ij}$
for each $1 \leqslant k, \ell \leqslant N$ do

$$
\begin{aligned}
d_{k\ell} = \text{minimum value of} \quad & \sum_{i,j=0}^{N} c_{ij} \cdot \lambda_{ij} \\
\text{subject to} \quad & \sum_{i=0}^{N} \lambda_{ij} = \pi(s_k, s_j), \quad 0 \leqslant j \leqslant N, \\
& \sum_{j=0}^{N} \lambda_{ij} = \pi(s_\ell, s_i), \quad 0 \leqslant i \leqslant N, \\
& 0 \leqslant \lambda_{ij}, \quad 0 \leqslant i, j \leqslant N.
\end{aligned}
$$

The presentation above is aimed at clarity. There is an obvious redundancy in that the distance matrix is always symmetric and has 0s along its diagonal, so we only ever need to calculate $d_{k\ell}$ for $k < \ell$.

9. Conclusion

At the time we came up with our algorithm, we only were aware of the Hutchinson metric. Only later we found out that the metric is known under many different names including the Kantorovich metric and the Wasserstein (Vaserstein or even Waserstein) metric. In the initial derivation of our algorithm, we showed that the characterization of d_{ϕ_n} in Theorem 19 can be reduced to a linear programming problem. Using the duality theorem of linear programming, we transformed the problem into the linear programming problem of Section 8. We refer the reader to [6] for more details. Here, we use the more general Kantorovich–Rubinstein duality theorem instead.

In this paper, we consider probabilistic transition systems without labels to simplify the presentation. However, all our results can easily be generalized to a setting with labels. The coalgebras of the functor

$$\mathcal{P}^L : \mathbb{CM}et_1 \to \mathbb{CM}et_1,$$

where L is the finite set of labels, represent probabilistic transition systems the transitions of which are labelled. To compute the d_{ϕ_n}-distance between states of a labelled system, for each label we consider only the transitions with that label and compute the d_{ϕ_n}-distance between the states, and take the maximum of all the computed distances.

Desharnais et al. [13] gave a decision procedure to calculate distances. Their decision procedure involves the generation of a representative set of formulae of their logic. They only consider formulae with a restricted number of nested occurrences of the modal connective. This corresponds to our approximation of $d_{fix(\Phi)}$ by d_{ϕ_n}. In both cases, the depth at which observations are considered is restricted. Their decision procedure calculates distances in exponential time—the main contribution of this paper is to give a strongly polynomial algorithm.

Recall that the functor $\mathcal{P}$ contains a scaling by a factor c smaller than 1. As we have shown in [4], a terminal $\mathcal{P}$-coalgebra and, hence, a behavioural pseudometric exists also in case that c equals 1. However, in that case we cannot exploit Proposition 12 to conclude that this behavioural pseudometric can be approximated to some prescribed degree of accuracy. Whether this behavioural pseudometric can be approximated is still an open problem.

Desharnais et al. [15] have shown that finite-state probabilistic transition systems are dense in the space of all probabilistic transition systems in the behavioural pseudometric that we have been considering in this paper. More precisely they provide an approximation construction that, given an infinite-state probabilistic transition system and a prescribed degree of accuracy ε, computes a finite-state approximation that is within distance ε of the original system (see proof of [12, Corollary 6.2.3]). One possible direction for future research would be to combine the approximation techniques (like, for example, the algorithm presented in [3]) with our algorithm to compute distances between states of systems with infinitely many states.

The network in the introduction is very similar to one occurring in the proof of the 'splitting lemma' by Jones and Plotkin in [23], except there one has a network whose edges have capacities rather than costs. In fact, hiding in Section 8 there is a splitting lemma for the metric on Borel probability measures which can be used to characterize $\mathcal{M}(X)$ as a free algebra (see [4] for some details). The same network also appears in the proof in [37] that the functor $\mathcal{P}'$ of de Vink and Rutten preserves weak pullbacks.

Our characterization of the behavioural pseudometric in terms of a linear programming problem has had impact on the definition of a pseudometric analogue of weak probabilistic bisimilarity by Desharnais et al. [14]. Also de Alfaro et al. [9,10], Deng et al. [11] and Gupta et al. [19] have defined behavioural pseudometrics along those lines.

Acknowledgements

We would like to thank the referees for their constructive feedback.

References

[1] P. America, J.J.M.M. Rutten, Solving reflexive domain equations in a category of complete metric spaces, J. Comput. System Sci. 39 (3) (1989) 343–375.

[2] S. Banach, Sur les opérations dans les ensembles abstraits et leurs applications aux équations intégrales, Fund. Math. 3 (1922) 133–181.

[3] A. Bouchard-Cote, N. Ferns, P. Panangaden, D. Precup, An approximation algorithm for labelled Markov processes towards realistic approximation, in: Proc. Second Internat. Conf. on the Quantitative Evaluation of Systems, Torino, September 2005, IEEE, New York, 2005.

[4] F. van Breugel, C. Hermida, M. Makkai, J. Worrell, An accessible approach to behavioural pseudometrics, in: L. Caires, G.F. Italiano, L. Monteiro, C. Palamidessi, M. Yung (Eds.), Proc. 32nd Internat. Colloq. on Automata Languages and Programming, Lecture Notes in Computer Science, Vol. 3580, Lisbon, July 2005, Springer, Berlin, 2005, pp. 1018–1030.

[5] F. van Breugel, J. Worrell, Towards quantitative verification of probabilistic transition systems, in: F. Orejas, P.G. Spirakis, J. van Leeuwen (Eds.), Proc. 28th Internat. Colloq. on Automata, Languages and Programming, Lecture Notes in Computer Science, Vol. 2076, Crete, July 2001, Springer, Berlin, 2001, pp. 421–432.

[6] F. van Breugel, J. Worrell, An algorithm for quantitative verification of probabilistic transition systems, in: K.G. Larsen, M. Nielsen (Eds.), Proc. 12th Internat. Conf. on Concurrency Theory, Lecture Notes in Computer Science, Vol. 2154, Aalborg, August 2001, Springer, Berlin, 2001, pp. 336–350.

[7] F. van Breugel, J. Worrell, A behavioural pseudometric for probabilistic transition systems, Theoret. Comput. Sci. 331 (1) (2005) 115–142.

[8] V. Chvátal, Linear Programming, W.H. Freeman and Company, New York, San Francisco, 1983.

[9] L. de Alfaro, Quantitative verification and control via the mu-calculus, in: R.M. Amadio, D. Lugiez (Eds.), Proc. 14th Internat. Conf. on Concurrency Theory, Lecture Notes in Computer Science, Vol. 2761, Marseille, September 2003, Springer, Berlin, 2003, pp. 102–126.

[10] L. de Alfaro, T.A. Henzinger, R. Majumdar, Discounting the future in systems theory, in: J.C.M. Baeten, J.K. Lenstra, J. Parrow, G.J. Woeginger (Eds.), Proc. 30th Internat. Colloq. on Automata, Languages and Programming, Lecture Notes in Computer Science, Vol. 2719, Eindhoven, June/July 2003, Springer, Berlin, 2003, pp. 1022–1037.

[11] Y. Deng, T. Chothia, C. Palamidessi, J. Pang, Metrics for action-labelled quantitative transition systems, in: Proc. Third Workshop on Quantitative Aspects of Programming Languages, Electronic Notes in Theoretical Computer Science, Edinburgh, April 2005, Elsevier, Amsterdam, 2005.

[12] J. Desharnais, Labelled Markov processes, Ph.D. Thesis, McGill University, Montreal, November 1999.

[13] J. Desharnais, V. Gupta, R. Jagadeesan, P. Panangaden, Metrics for labeled Markov systems, in: J.C.M. Baeten, S. Mauw (Eds.), Proc. 10th Internat. Conf. on Concurrency Theory, Lecture Notes in Computer Science, Vol. 1664, Eindhoven, August 1999, Springer, Berlin, 1999, pp. 258–273.

[14] J. Desharnais, V. Gupta, R. Jagadeesan, P. Panangaden, The metric analogue of weak bisimulation for probabilistic processes, in: Proc. 17th Annual IEEE Symp. on Logic in Computer Science, Copenhagen, July 2002, IEEE, New York, 2002, pp. 413–422.

[15] J. Desharnais, V. Gupta, R. Jagadeesan, P. Panangaden, Metrics for labeled Markov systems, Theoret. Comput. Sci. 318 (3) (2004) 323–354.

[16] G.A. Edgar, Integral, Probability, and Fractal Measures, Springer, New York, 1998.

[17] A. Giacalone, C.-C. Jou, S.A. Smolka, Algebraic reasoning for probabilistic concurrent systems, in: Proc. IFIP WG 2.2/2.3 Working Conf. on Programming Concepts and Methods, Sea of Gallilee, April 1990, North-Holland, Amsterdam, 1990, pp. 443–458.

[18] M. Giry, A categorical approach to probability theory, in: B. Banaschewski (Ed.), Proc. Internat. Conf. on Categorical Aspects of Topology and Analysis, Lecture Notes in Mathematics, Vol. 915, Ottawa, August 1981, Springer, Berlin, 1981, pp. 68–85.

[19] V. Gupta, R. Jagadeesan, P. Panangaden, Approximate reasoning for real-time probabilistic processes, in: Proc. First Internat. Conf. on the Quantitative Evaluation of Systems, Enschede, September 2004, IEEE, Amsterdam, 2004, pp. 304–313.

[20] M. Hennessy, R. Milner, Algebraic laws for nondeterminism and concurrency, J. ACM 32 (1) (1985) 137–161.

[21] J.E. Hutchinson, Fractals and self similarity, Indiana University Mathematics Journal 30 (5) (1981) 713–747.

[22] B. Jacobs, J. Rutten, A tutorial on (co)algebras and (co)induction, Bull. EATCS 62 (1997) 222–259.

[23] C. Jones, G.D. Plotkin, A probabilistic powerdomain of evaluations, in: Proc. Fourth Annu. IEEE Symp. on Logic in Computer Science, Pacific Grove, California, June 1989, IEEE, Amsterdam, 1989, pp. 186–195.

[24] L.V. Kantorovich, On the transfer of masses, Dokl. Akad. Nauk 37 (2) (1942) 227–229 (in Russian); Translated in Management Sci. 5(1) (1958) 1–4.

[25] L.V. Kantorovich, G.Sh. Rubinstein, On the space of completely additive functions, Vestnik Leningrad. Univ. 3 (2) (1958) 52–59 (in Russian).

[26] J. Lambek, Subequalizers, Canadian Mathematical Bulletin 13 (3) (1970) 337–349.

[27] K.G. Larsen, A. Skou, Bisimulation through probabilistic testing, Inform. Comput. 94 (1) (1991) 1–28.

[28] E.J. McShane, Extension of range of functions, Bull. Amer. Math. Soc. 40 (12) (1934) 837–842.

[29] R. Milner, A Calculus of Communicating Systems, Springer, Berlin, 1980.
[30] J. Orlin, A faster strongly polynomial minimum cost flow algorithm, in: Proc. 20th Annu. ACM Symp. on Theory of Computing, Chicago, May 1988, ACM, New York, 1988, pp. 377–387.
[31] D. Park, Concurrency and automata on infinite sequences, in: P. Deussen (Ed.), Proc. Fifth GI-Conf. on Theoretical Computer Science, Lecture Notes in Computer Science, Vol. 104, Karlsruhe, March 1981, Springer, Berlin, 1981, pp. 167–183.
[32] K.R. Parthasarathy, Probability Measures on Metric Spaces, Academic Press, New York, 1967.
[33] J.J.M.M. Rutten, Universal coalgebra: a theory of systems, Theoret. Comput. Sci. 249 (1) (2000) 3–80.
[34] E. Tardos, A strongly polynomial minimum cost circulation algorithm, Combinatorica 5 (3) (1985) 247–255.
[35] D. Turi, J.J.M.M. Rutten, On the foundations of final semantics: non-standard sets, metric spaces, partial orders, Math. Struct. Comput. Sci. 8 (5) (1998) 481–540.
[36] L. Vaserstein, Markov processes over denumerable products of spaces, describing large systems of automata, Problems Inform. Transmission 5 (3) (1969) 47–52.
[37] E.P. de Vink, J.J.M.M. Rutten, Bisimulation for probabilistic transition systems: a coalgebraic approach, Theoret. Comput. Sci. 221 (1/2) (1999) 271–293.
[38] H. Whitney, Analytic extensions of differentiable functions defined in closed sets, Trans. Amer. Math. Soc. 36 (1) (1934) 63–89.